

2023년 K-iium 의료인공지능경진대회

AI-BASED DIAGNOSIS OF ANEURYSMS

주최 경북대학교병원, 부산대학교병원, 전남대학교병원

팀명 키움 화이팅

팀원 김수인 곽엘림 박민영

박지은 하민세 한규원



GOAL

프로젝트 목표

GOAL

- 익명화 의료 데이터 셋 활용 인공지능(AI) SW 개발
- 익명화된 뇌혈관조형술 데이터 이용
- 뇌동맥류 여부, 위치를 진단

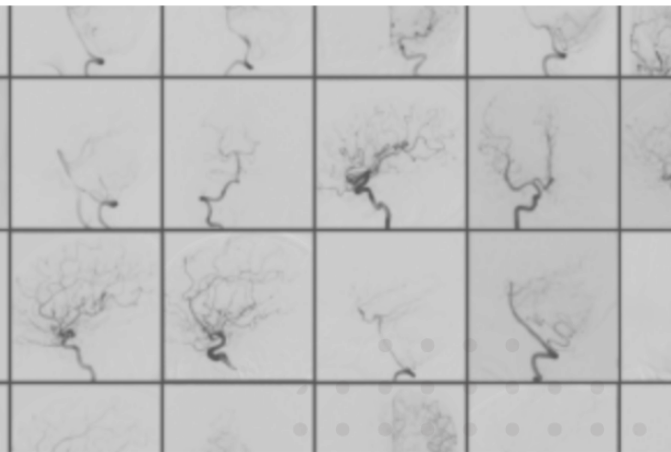
DATA

- 총 1127명의 환자 뇌혈관조형술 영상
- 환자당 좌/우 X 내경/추골 X 전면/후면 영상 포함
- 환자 1명 당 총 8종의 영상

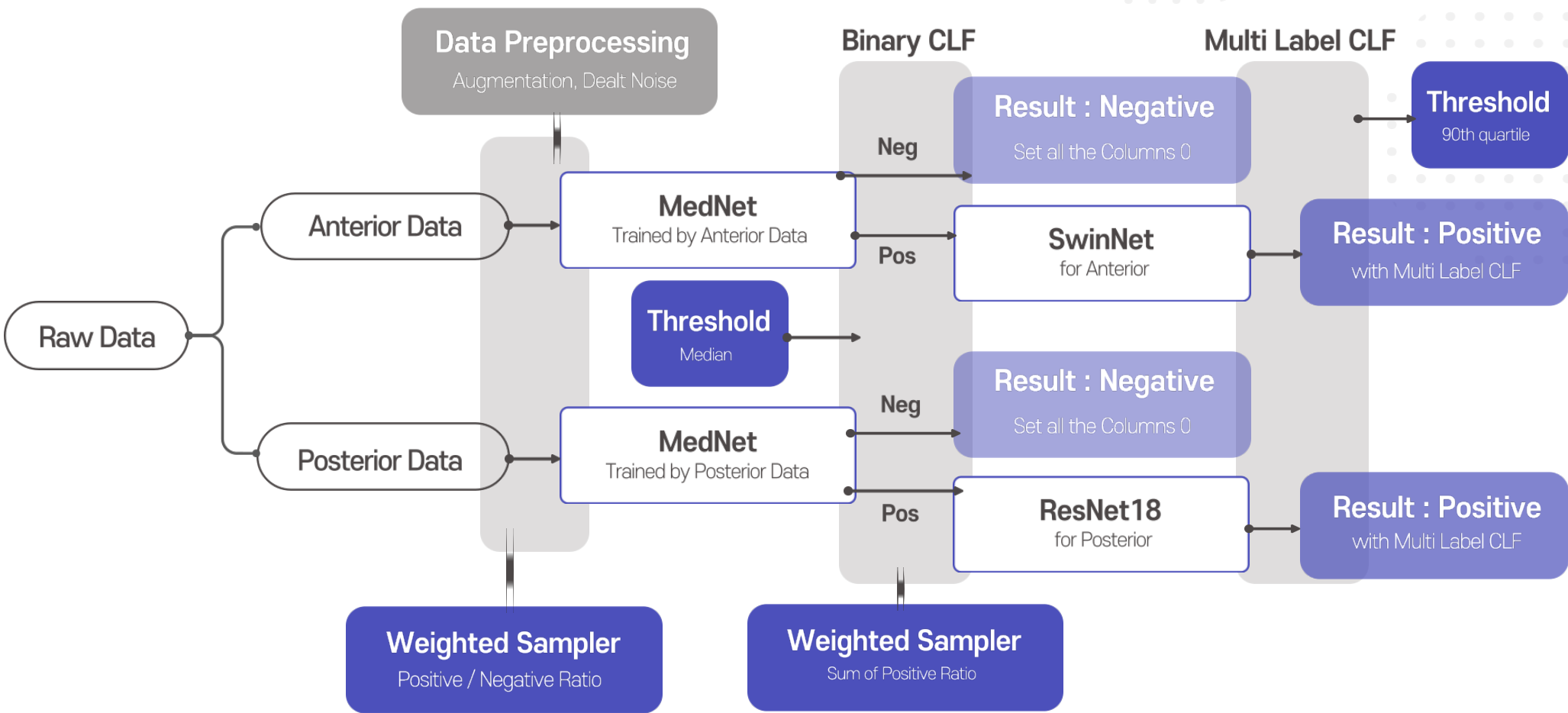


DATA

데이터 구성



INFERENCE PIPELINE 전체 추론 파이프라인



Data Splitting

분류기 학습을 위한 데이터 분할

Original Data

	Index	Aneurysm	L_ICA	R_ICA	L_PCOM	R_PCOM	L_AntChor	R_AntChor	L_ACA	R_ACA	...	L_VA	R_VA	L_PICA	R_PICA	L_SCA	R_SCA	BA	L_PCA	R_PCA
0	1001	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0
1	1002	1	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0
2	1004	1	0	0	0	1	0	0	0	0	...	0	0	0	0	0	0	0	0	0
3	1005	1	0	0	0	1	0	1	0	0	...	0	0	0	0	0	1	0	0	0
4	1006	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0
...
1122	2607	1	0	0	0	0	0	0	0	0	...	0	1	0	0	0	0	0	0	0
1123	2608	1	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0
1124	2609	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0
1125	2610	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0
1126	2611	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0

For Binary CLF

Anterior

	Index	sum	path
4	1002	1	/content/drive/MyDrive/2023_k_iu_m_composition/...
5	1002	1	/content/drive/MyDrive/2023_k_iu_m_composition/...

Posterior

	Index	sum	path
4	1002	0	/content/drive/MyDrive/2023_k_iu_m_composition/...
5	1002	0	/content/drive/MyDrive/2023_k_iu_m_composition/...

For Multi-Label CLF

Anterior

	Index	ICA	AntChor	ACA	ACOM	NCA	sum	path
4	1002	0	0	0	1	0	1	/content/drive/MyDrive/2023_k_iu_m_composition/...
5	1002	0	0	0	1	0	1	/content/drive/MyDrive/2023_k_iu_m_composition/...

Posterior

	Index	VA	PICA	SCA	BA	PCA	PCOM	sum	path
4	1002	0	0	0	0	0	0	0	/content/drive/MyDrive/2023_k_iu_m_composition/...
5	1002	0	0	0	0	0	0	0	/content/drive/MyDrive/2023_k_iu_m_composition/...

Data Preprocessing

학습성능 향상을 위한 데이터 전처리

01 Resize/Normalization

- 계산 효율성 증대
- 학습과 추론에 일관된 결과 획득
- 학습 수행 일반화

02

- 정상 이미지는 Resize + 정규화만 진행
- 이미지 가장자리의 여백은 제거 후 Resize
- 이미지 내 텍스트 있는 경우 배경 픽셀값으로 변환

03 Augmentation

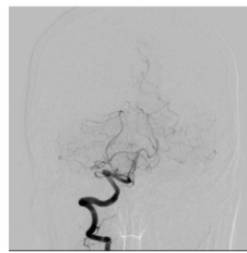
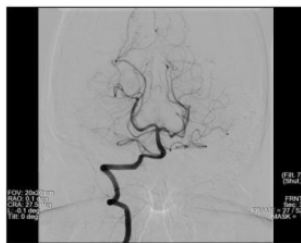
Median Blur : 이미지의 노이즈 감소 -> **사용**

Bitwise_not : 색상 반전에 이용 -> **사용**

Horizontal Flip : 좌우반전, 성능감소로 -> **미사용**

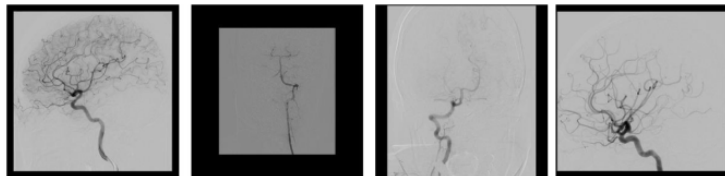
Vertical Flip : 상하반전, 성능감소로 -> **미사용**

텍스트가 있는 이미지

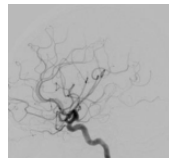
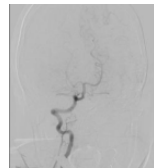
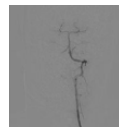
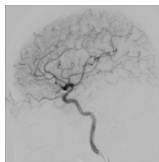


해당 이미지는 같은 위치의 사진이 아닙니다

여백이 있는 이미지



여백 삭제 후 Resize



Model / Loss

성능 최우수 모델 및 손실함수

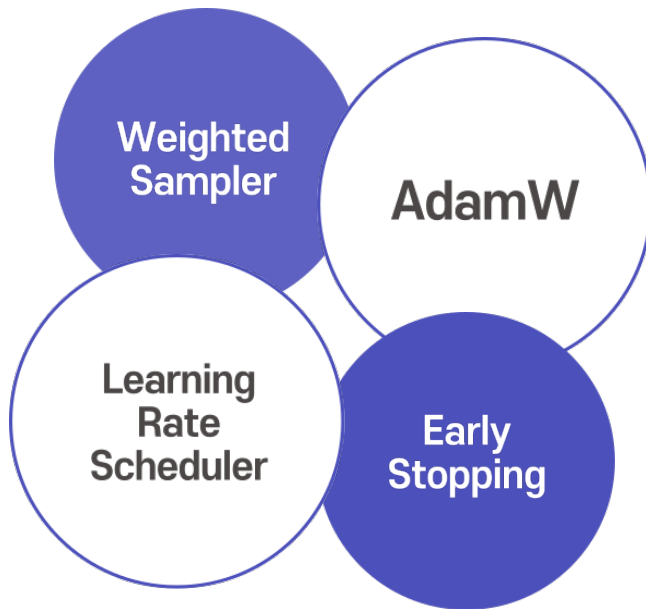
Task	Data for Training	Model	Loss	Optimizer
Binary CLF	Anterior Dataset	MedNet(ResNet10)	BCE Loss	AdamW
Binary CLF	Posterior Dataset	MedNet(ResNet10)	BCE Loss	AdamW
Multi-Label CLF	Anterior Dataset	SwinNet	Asymmetric Loss	AdamW
Multi-Label CLF	Posterior Dataset	ResNet18	Asymmetric Loss	AdamW

Metric : ROC(Binary) / F1Score(Multi-Label)

Tuning Methods

성능 향상을 위한 튜닝 Methods

- Data Imbalancing 문제 해결을 위해 Binary CLF에선 양/음성 데이터 비율로, Multi-Label CLF에선 한 데이터 내 양성 판정 위치 합 비율에 따라 Weighted Sampling을 수행



- 고정 Learning Rate로 학습 시 Loss/Val Score 수렴 X -> Scheduler 도입으로 Learning Rate 1/10 축소해 수렴 유도 (patience = 2)

- 가중치 감소 완화에 효과적인 Adam의 변종 AdamW를 Optimizer로 도입해 학습 속도와 성능 개선
- 극심한 데이터 불균형으로 인해 10~20 Epoch 이후 모두 Frequent Data에 Overfitting 되는 문제 해결을 위해 Early Stopping 도입 : 5회 이상 진전이 없으면 학습 종료

Results & Improvements

최종 결과 및 추후 보완점

Results

Task	Data for Training	Train Loss	Val Loss	Val Score
Binary CLF	Anterior Dataset	0.07472	2.17023	0.60
Binary CLF	Posterior Dataset	0.06718	0.45695	0.55
Multi-Label CLF	Anterior Dataset	16.89743	17.43649	0.51
Multi-Label CLF	Posterior Dataset	6.32751	13.99973	0.38

Metric : ROC(Binary) / F1Score(Multi-Label)

Improvements

1. Improve Metric Reliability
2. Train with more Data
3. Use Segmentation to Train Data
4. Use XAI to explain the Inference Result
5. Use more Complex/Massive Model

Sources

출처

Asymmetric Loss

https://github.com/Alibaba-MIL/ASL/blob/main/src/loss_functions/losses.py

MedNet Weight

<https://huggingface.co/TencentMedicalNet/MedicalNet-Resnet18>

2023 K-ium 의료인공지능 경진대회

**THANK
YOU**